AD-A119 471 RAND CORP SANTA MONICA CA F/G 12/1
THE RELATION AMONG THE LIKEHOOD RATIO-, WALD-, AND LAGRANGE MUL--ETC(U)
APR 82 D F KOHLER
UNCLASSIFIED RAND/P-6756 NL

1 of 1

END
DATE
FILMED
10 .82
DTIC

AD A119471

THE RELATION AMONG THE LIKELIHOOD RATIO-, WALD-, AND LAGRANGE MULTIPLIER TESTS AND THEIR APPLICABILITY TO SMALL SAMPLES

Daniel F. Kohler

April 1982

P-6756

1. INTRODUCTION

The Lagrange multiplier test (LM), the likelihood ratio test (LR), and the Wald test (W), are frequently proposed as alternative means for testing parametric restrictions in a linear regression model. Since they all converge to the same limiting Chi-square distribution they have asymptotically the same power characteristics. Consequently the three tests are considered equivalent alternatives for large samples, and the choice among them is usually made on convenience grounds (e.g., ease of estimation).

However, Savin (1976) and Berndt and Savin (1977), among others, have shown that the numerical value of the LM test statistic is always less than that of LR, which in turn is less than that of W. This implies that if we use the same critical value, as seems indicated by the fact that all three tests converge to the same limiting Chi-square distribution, the Wald test will reject the null hypothesis most often, while the LM test will reject it least often. This difference in rejection probabilities raises the possibility of conflicting conclusions from the three tests.

In this paper we show that the three tests are monotonic functions of each other. This implies that they have the same power characteristics. If the critical values are specified such as to equate the probabilities of type I error, the probabilities of type II error will be equal as well, and conflicting results are impossible.

The critical values of the three tests are related to each other by the same monotonic functions as the three test statistics. The "conflicts" pointed out in the literature arise only when the exact

critical values for each test are replaced by the asymptotically justified critical value of a Chi-square distribution. The extent to which this critical value differs from the exact one varies from test to test.

Determining the exact critical values for finite samples is somewhat difficult, because the exact finite sample distributions of the LM and LR test are generally not known. However, under the assumption of normality, the W test can be transformed into an F-test by applying the standard degrees of freedom correction. Furthermore, since the LR and LM tests are functionally related to the W test, they can be transformed into the same test statistic. We can thus conduct an exact finite sample test, based either on the LM, LR or W statistic without running the risk of obtaining conflicting results.

2. THE FORM OF THE TESTS

Consider the model:

$$Y = X\beta + \varepsilon \tag{1}$$

$$\varepsilon \sim N(0,\sigma^2 I) \tag{2}$$

where X is of rank k.

Let $\hat{\beta}$ and $\hat{\sigma}^2$ be the maximum-likelihood estimators obtained by unconstrained maximization of the likelihood function, and $\tilde{\beta}$ and $\tilde{\sigma}^2$ the estimators obtained with the constraint $r = R\tilde{\beta}$ imposed. Furthermore, we shall need an estimator for the Lagrange multiplier ($\hat{\mu}$) and the ratio of the unconstrained to constrained maxima of the likelihood function (λ). The three test statistics can be written as:

$$LR = -2 \log \lambda \tag{3}$$

$$W = [(r - R\hat{\beta})'(R(X'X)^{-1}R')^{-1}(r - R\hat{\beta})]/\hat{\sigma}^2 \tag{4}$$

$$LM = [\hat{\mu}(R(X'X)^{-1}R')\hat{\mu}]/\tilde{\sigma}^2 \tag{5}$$

where R is the matrix of constraints and is of rank q.

To simplify the notation let $A = R(1/n\ X'X)^{-1}R'$. From the first-order conditions for maximizing the likelihood function subject to the constraint we can obtain an expression for $\hat{\mu}$:

$$\hat{\mu} = (R(X'X)^{-1}R')^{-1}(r - R\hat{\beta}) \quad . \tag{6}$$

We can now rewrite Eqs. (4) and (5) as

$$W = n[(r - R\hat{\beta})'A^{-1}(r - R\hat{\beta})]/\hat{\sigma}^2 \tag{7}$$

$$LM = n[(r - R\hat{\beta})'A^{-1}(r - R\hat{\beta})]/\tilde{\sigma}^2 \quad . \tag{8}$$

This implies that we can express the LM test as

$$LM = \frac{\hat{\sigma}^2}{\tilde{\sigma}^2} W \quad . \tag{9}$$

If we make use of the fact that the likelihood ratio $\lambda = (\hat{\sigma}^2/\tilde{\sigma}^2)^{n/2}$ we can express Eq. (9) as

$$LM = \lambda^{2/n} \cdot W \quad . \tag{10}$$

Equation (10) makes it obvious that the LM and W test statistics are asymptotically equivalent.

To express LR as a function of W consider $\tilde{\sigma}^2 \equiv (Y - X\tilde{\beta})'(Y - X\tilde{\beta})/n$ and $\tilde{\beta} = \hat{\beta} + n(X'X)^{-1}R'A^{-1}(r - R\hat{\beta})$. Then, letting $e = Y - X\hat{\beta}$, we get by substitution:

$$\begin{aligned}\tilde{\sigma}^2 &= 1/n\{e'e - 2e'X(1/nX'X)^{-1}R'A^{-1}(r - R\hat{\beta}) \\ &\quad + (r - R\hat{\beta})'A^{-1}n[R(X'X)^{-1}X'X(X'X)^{-1}R']A^{-1}(r - R\hat{\beta})\} \\ &= \hat{\sigma}^2 + (r - R\hat{\beta})'A^{-1}(r - R\hat{\beta}) \quad .\end{aligned} \tag{11}$$

We can, therefore, express Eq. (7) as

$$W = n(\tilde{\sigma}^2 - \hat{\sigma}^2)/\hat{\sigma}^2 = n(\tilde{\sigma}^2/\hat{\sigma}^2 - 1) \tag{12}$$

which is equivalent to

(13) $$W = n(\lambda^{-2/n} - 1) \quad .$$

Combining Eqs. (13) and (10) leads to

(14) $$LM = n(1 - \lambda^{2/n}) \quad .$$

Let $\theta \equiv \lambda^{-2/n}$ which is larger or equal to unity with probability one. Then, from Eqs. (3), (13), and (14) the three tests can be rewritten as

(15) $$LR = n \log \theta$$

(16) $$W = n(\theta - 1)$$

(17) $$LM = n(1 - 1/\theta) \quad .$$

Thus, since $(1 - 1/\theta) < \log \theta < \theta - 1$ we have the well-known inequality relationship

(18) $$LM \leq LR \leq W \quad .$$

3. THE THREE TESTS AS TRANSFORMATIONS OF THE F-TEST*

If we divide Eq. (4) by q we get

$$(19) \qquad W/q = [(r - R\hat{\beta})'(R(X'X)^{-1}R')^{-1}(r - R\hat{\beta})]/\hat{\sigma}^2 \cdot q$$

which is distributed as $F(q,n - k)$. From Eqs. (15) and (16) we can express W/q as a function of LR, i.e.,

$$(20) \qquad W/q = n \cdot [\exp(LR/n) - 1]/q$$

which also follows an $F(q,n - k)$ distribution. With Eqs. (19) and (20) we can thus transform either W or LR into an F statistic of which we know the exact finite sample distribution. This allows us to conduct exact finite sample tests.

Much of the justification of the LM test is based on the fact that it relies exclusively on parameter estimates from the restricted model. That makes it particularly attractive in situations where the unrestricted model is difficult or impossible to estimate. In the previous section we have given a formula (Eq. (9)) by which we could compute the W test statistic from the LM statistic. But it relies on the unconstrained estimate $\hat{\sigma}^2$ which defeats the major advantage of the LM test. We have to find an alternative way of calculating an F statistic from the LM statistic.

Recall that the LM test is based on maximizing the likelihood function, subject to the constraint, by the Lagrange method and then

*Throughout this section we assume that the F-test is defined. In particular we require that $n - k > 0$.

testing how close the estimated Lagrange multipliers ($\hat{\mu}$) are to zero. We have shown that under the assumptions of the linear model this is equivalent to testing

$$(21) \qquad LM = n[(r - R\hat{\beta})'A^{-1}(r - R\hat{\beta})]/\tilde{\sigma}^2 \quad .$$

But note that we can express $\tilde{\sigma}^2$ as a function of the unconstrained parameters. In particular we have (see Eq. (11))

$$\tilde{\sigma}^2 = \hat{\sigma}^2 + (r - R\hat{\beta})'A^{-1}(r - R\hat{\beta})$$

so that we can express the LM statistic as

$$(22) \qquad LM = n \cdot [(r - R\hat{\beta})'A^{-1}(r - R\hat{\beta})]/[\hat{\sigma}^2 + (r - R\hat{\beta})'A^{-1}(r - R\hat{\beta})] \quad .$$

Dividing LM by n and inverting results in

$$(23) \qquad \frac{n}{LM} = 1 + \frac{\hat{\sigma}^2}{(r - R\hat{\beta})'A^{-1}(r - R\hat{\beta})}$$

which can be transformed into

$$(24) \qquad \left[\frac{n}{LM} - 1\right]^{-1} = \frac{(r - R\ )'A^{-1}(r - R\hat{\beta})}{\hat{\sigma}^2} = \frac{W}{n} \quad .$$

For finite samples we can use the statistic

$$\frac{n}{q}\left[\frac{n}{LM} - 1\right]^{-1} = \frac{[(r - R\hat{\beta})'(R(X'X)^{-1}R')^{-1}(r - R\hat{\beta})]}{\hat{\sigma}^2 \cdot q} = \frac{W}{q}$$

which follows an F distribution with (q, n - k) degrees of freedom. Like the W and LR test statistics, the LM statistic can thus be

adjusted for sample size and degrees of freedom to result in a test statistic applicable for finite samples. As it does not require estimating the unconstrained model, it is particularly useful in situations where only the constrained version can readily be estimated.

4. CONCLUSIONS

An intuitively appealing analogy for the relationship between the three tests can be drawn from trigonometry. Testing linear restrictions can be compared to evaluating the angle α in Fig. 1.

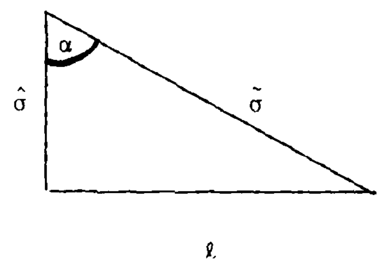

Fig. 1

The length ℓ is equal to the square root of $(r - R\hat{\beta})'(R(X'X)^{-1}R')^{-1}(r - R\hat{\beta})$. The W test statistic corresponds to $n \cdot [\text{tangent }(\alpha)]^2$, and the LM statistic to $n \cdot [\text{sine }(\alpha)]^2$. Since in a triangle the sine of α is always less than the tangent of α, this is yet another way of demonstrating LM < W.

The likelihood ratio λ corresponds to $\{[\text{cosine }(\alpha)]^{n/2}\}^2$. Accordingly, the LR test statistic is $-n \cdot \log[\text{cosine }(\alpha)]^2$. We can now use the basic trigonometric identities to express any one of the tests as a function of the other two. In particular, we know that

---

*Note that this quantity can also be calculated from the restricted regression estimate of $\hat{\mu}$ (see Eq. (6)).

tangent (α) = sine (α)/cosine (α) which, if raised to the power two, corresponds to Eq. (10).

Continuing with the trigonometric analogy, we see that Eq. (11) corresponds to the theorem of Pythagoras. Thus, if we have an estimate of any two of the three sides of the triangle in Fig. 1, either from the unconstrained regression ($\hat{\sigma},\hat{\beta}$) or from the constrained one ($\tilde{\sigma},\hat{\mu}$), we can calculate the third side and then conduct any one of the three tests.

Note that $\hat{\sigma}$ and (r - $R'\hat{c}$) are orthogonal to each other. In statistical terms this means that they are independently distributed. We can thus compute the exact small sample distribution of $n \cdot [\text{tangent } (\alpha)]^2$, which turns out to be a multiple of an F distribution. For the other two tests this is not directly possible, because the components going into the calculation of sine (α) and cosine (α) are not independent of each other. However, by expressing these tests as deterministic functions of tangent (α), e.g., through the trigonometric relations described above, we can obtain test statistics with known finite sample distributions as well.

References

Aitchison, J. and Silvey, S. C. (1958), "Maximum-Likelihood Estimation of Parameters Subject to Restraints," Annals of Mathematical Statistics, 29, 813-828

Berndt, E. and Savin, N. E. (1977), "Conflict Among Criteria for Testing Hypothesis in the Multivariate Linear Regression Model," Econometrica, 45, 1263-1278

Breusch, T. S. (1979), "Conflict Among Criteria for Testing Hypothesis: Extension and Comments," Econometrica, 47, 203-207

Breusch, T. S. and Pagan, A. R. (1980), "The Lagrange Multiplier Test and Its Application to Model Specification in Economics," Review of Economic Studies, XLVII, 239-253

Hardy, G. H., Littlewood, J. E. and Polya, G. (1967), Inequalities, Cambridge University Press

Kendall, M. G. and Stuart, A. (1973), The Advanced Theory of Statistics, Vol. 2, 3rd edition, Hafner, New York

Kohler, D. F., "Conflict Among Testing Procedures in a Linear Regression Model," Wayne Economic Papers No. 81, Wayne State University, Detroit, 1979

Pearson, C. E. (ed) (1974), Handbook of Applied Mathematics, Van Nostrand Reinhold Co., New York

Savin, N. E. (1976), "Conflict Among Testing Procedures in a Linear Regression Model With Autoregressive Disturbances," Econometrica, 44, 1303-1315

Silvey, S. D. (1959), "The Lagrange Multiplier Test," Annals of Mathematical Statistics, 30, 389-407

Wald, A. (1943), "Tests of Hypotheses Concerning Several Parameters When the Number of Observations is Large," Transactions of the American Mathematical Society, 54, 476-482

DAT
FILM
0-8